Text-Based Active Authentication for Mobile Devices

Hataichanok Saevanee, Nathan Clarke, Steven Furnell, Valerio Biscione

Text-Based Active Authentication for Mobile Devices

Hataichanok Saevanee[1], Nathan Clarke[1,3], Steven Furnell[1,3] and Valerio Biscione[2]

[1]Centre for Security, Communications and Network Research,
[2]Centre for Robotics and Neural Systems,
Plymouth University, Plymouth, United Kingdom
[3]Security Research Institute, Edith Cowan University,
Perth, Western Australia
info@cscan.org

Abstract. As modern mobile devices are increasing in their capability and accessibility, they introduce additional demands in terms of security – particularly authentication. With the widely documented poor use of PINs, Active Authentication is designed to overcome the fundamental issue of usable and secure authentication through utilizing biometric-based techniques to continuously verify user identity. This paper proposes a novel text-based multimodal biometric approach utilizing linguistic analysis, keystroke dynamics and behavioral profiling. Experimental investigations show that users can be discriminated via their text-based entry, with an average Equal Error Rate (EER) of 3.3%. Based on these findings, a framework that is able to provide robust, continuous and transparent authentication is proposed. The framework is evaluated to examine the effectiveness of providing security and user convenience. The result showed that the framework is able to provide a 91% reduction in the number of intrusive authentication requests required for high security applications.

Keywords: Active authentication ·Transparent authentication · Continuous authentication · Multimodal · Biometric · Mobile devices

1 Introduction

Mobile devices are commonplace with over 6 billion subscribers worldwide [1]. With the rapid development of mobile network technology and the increasing popularity of mobile devices, modern devices are capable of providing a wide range of services and applications over multiple networks. The plethora of functionalities offered by the mobile device enables users to store increasing amounts of a wider variety of information from business to personal and sensitive data. A series of studies have highlighted the potential risk of mobile device misuse through the storing of personal information (e.g. home address), security credentials (e.g. PIN codes, user names and passwords) and business data (e.g. customer data) [2,3].

Although PIN or password authentication is available on most mobile devices, a survey conducted by [4] demonstrated that a third of mobile users do not protect their devices with this simple technique. Furthermore, the poor use of PIN or pass-

word techniques when they are used is also widely documented in several studies [4,5]. A fundamental weakness of the PIN is that as a point-of-entry approach, once the user has been successfully authenticated, they obtain access to the system without having to re-authenticate. Several studies [6,7] proposed Active Authentication or transparent authentication to overcome the fundamental issue and more closely associate the authentication and access control decisions. There are a number of biometric techniques that have the potential to be used for authentication in a transparent and thus continuous fashion, such as keystroke dynamics, behavioral profiling, gait recognition, speaker verification and facial recognition. Unfortunately, research has demonstrated that using a single biometric may be inadequate for verification due to a variety of reasons, such as noise in the sample data, the unavailability of a sample at a given time and the underlying performance of the technique [8]. To overcome this limitation within traditional the point-of-entry domain, several researchers have proposed the use of multiple biometric modalities, which have demonstrated increased accuracy of verification [9,10,11].

This paper presents the findings of a research study exploring the application of multimodal biometric authentication in a transparent fashion to text-based entry. As users frequently use their mobile device to send SMS text messages (over 9.8 trillion in 2012), social network posts, emails and tweets, it was felt this medium provided a frequent opportunity to capture samples [12]. The focus upon text-based entry provides the possibility to apply keystroke dynamics, linguistic analysis and behavioral profiling. It is the aim of this paper to present the results of an exhaustive investigation into optimizing the recognition performance and an evaluation of the security processes required to maximize the security of the approach whilst minimizing user inconvenience. Section 2 presents the state of the art in behavioral biometrics that have been applied in the mobile domain. Section 3 describes the feasibility study of multimodal biometric. Based upon the results, a novel text-based multimodal framework that will provide the verification of a mobile user's identity in a continuous and transparent manner is proposed in Section 4 and then evaluated through simulation in Section 5. The paper concludes by highlighting the future direction of research in Section 6.

2 Text-based behavioral biometric for mobile devices

With the rapid evolution of mobile devices, utilizing biometrics on them has become a reality. Many mobile devices come equipped with a number of hardware components that are able to be used for capturing a variety of biometric traits, enabling several biometric approaches to be deployed – such as keystroke dynamics, behavioral profiling and voice recognition. For example, Apple has now incorporated TouchID, a fingerprint-based approach, and Google has Face Unlock for its Android Operating System [13,14]. To date, however, these are point-of-entry solutions that focus upon usability rather than security. Of interest in this research is the use of three behavioral biometric techniques: linguistic profiling, keystroke dynamics and behavioral profiling. It is hypothesized that the integration of these three techniques together offers the opportunity

to improve upon the usability through transparent capture, improve the overall recognition performance and mitigate the unavailability of samples at a given time.

Linguistic profiling is a behavioral biometric that identifies people based upon linguistic morphology. Previous studies have investigated the feasibility of linguistic profiling for several tasks such as text categorization, authorship identification and authorship verification. In the authorship verification domain, examples of writing from a single author are given to the system, which is then asked to confirm if the given texts were written by this author. According to previous studies [15], almost 1000 writing styles have been analyzed and both statistical and machine learning methods were used in the analytical process. Many studies have confirmed the good discriminating capability of linguistic features. Through using a machine learning method, the performance accuracies were in the range of 80%-100% [16,17]. However, there is no agreement on a best set of features for authorship verification and historically large volumes of text are required for the training dataset. The performance of linguistic profiling technique highly depends upon the combination of the selected features and classification models utilized.

Behavioral profiling aims to identify users based upon the way in which they interact with the services on their mobile device. Previous behavior-based studies have mainly focused upon the area of fraud detection. Research in mobile IDSs can be divided into two categories: call-based and mobility-based mechanisms. The former monitors user's calling behavior (e.g. start date of call and dial telephone number) that have been collected over a service provider's network during a period of time [18,19]. Based upon the theory that people have a predictable travelling pattern when they travel from one location to another, the mobility-based approach monitors a mobile user's location activities to detect abnormal behavior [20]. Through monitoring a user's calling or location activities, behavioral-based IDS can offer a high detection rate and ability to detect unforeseen attacks [18,19,20,21]. Depending upon application types, profiling techniques and classification approach, a study by [7] showed that behavioral profiling could be used for authentication on mobile devices with accuracies of between 87% and 98%.

Keystroke dynamics identifies a user based upon the typing pattern of a user, looking at characteristics of their interaction with a keyboard. Based upon previous studies, two main characteristics were identified: inter-key and hold time [24]. The inter-key is the duration between two successive keys. The hold-time represents the duration between the press down and releasing of a single key. Many studies have shown it is feasible to authenticate users successfully based upon usernames and passwords (i.e. in parallel with a typical Windows login request), with a commercial product on the market utilizing this technology [22, 23]. More recent studies [6, 24] investigated the possibility of using keystroke dynamics on mobile devices, showing the possibility of keystroke dynamic based authentication can be deployed in practice to provide an extra layer of security for mobile devices with an average accuracy of 87%.

Based upon the prior-art, these three techniques provide valuable discriminative information to permit identity authentication. All of the biometric traits of these three techniques can be captured during user interactions with a mobile device without a user explicit interaction to authenticate. In addition, no additional hardware is required to deploy these techniques. As a result, these approaches arguably provide a

cost effective and a non-intrusive solution for mobile handset authentication. Furthermore, a significant amount of prior research within the point-of-entry authentication domain [9,10,11] has concluded that using multiple biometric modalities can improve accuracy and reliability of single-modal systems. For example, using combination of fingerprint and face modality can achieve better performance than using single biometric, improving the accuracy of 2.3% at 0.1% FAR [25].

3 A feasibility study of text-based multimodal biometrics

Since no multimodal database availability where the above three biometric modalities are measured within the same individual, a standard practice employed within multi-biometrics is to combine the modalities from different datasets and create a virtual person [11]. The SMS corpus collected by the authors, a public mobile usage dataset provided by [26] and keystroke dataset provided by [24] were used in this experiment. An individual user from the linguistic profiling database was associated with an individual of keystroke and behavioral profiling database to create a virtual subject. As a result, a final database consisting of 30 users, each user having their SMS messages, keystroke and text messaging activity data was created and utilized in this experiment.

3.1 Experiment procedure

The experiments investigated the performance both of the individual techniques and their combination. To investigate the linguistic profiling's effectiveness; four types of linguistic features were examined: word profiling, lexical, syntactic and structural. The frequency distribution of a total 133 abbreviations and emotional words were used to create a user's word profiles, including 64 discriminating characteristics of every possible type of feature. To create a user profile, the t-test ranking measure was utilized to rank input features according to its discriminative capability. From the ranking list, features with a p value less than 0.05 were selected to create input vectors. The key to utilizing the t-test was to ensure a set of features that was as unique to the individual authorized user in comparison to the wider population. Therefore, the number of linguistic features required for discrimination will vary between users. Three different classification techniques: K-Nearest Neighbor (K-NN), the Radial Basis function (RBF) and Feed-Forward Multi-Layered Perceptron (FF-MLP) neural networks were utilized with differing network configurations - looking to optimum performance.

In the keystroke dynamics experiment, the hold time vector constructed from five letters: E, T, A, O and N were extracted. A number of analyses were undertaken using the FF-MLP neural network as it had demonstrated the better performance in previous studies over other techniques [24].

For the behavioral profiling technique, the following features were extracted: receiver's telephone number and location of texting. A number of analyses were undertaken, using a Radial Basis Function (RBF) neural network as it had performed the best in the prior study [7].

Table 1. Final dataset used in the experiments

	Training size	Testing size
Linguistic profiling	316	171
Keystroke dynamics	3339	171
Behavior profiling	1178	171

To perform the classification for the individual techniques, the dataset was divided into two groups: 171 data samples were used for the testing set and the remainder was used for training (as illustrated in Table 1). The pattern classification test was performed with one user acting as the valid user, while all others are acting as impostors (a standard procedure in this type of test) [6-8]. The Equal Error Rate (EER) was calculated to evaluate the system. The EER is the value where False Acceptance Rate (FAR) crosses the False Rejection Rate (FRR), and is typically used as a comparative measure within the biometric industry [28].

The multimodal experiment was conducted using all possible combination of three techniques. The results of each technique were combined at the matching-level - as each technique utilized different classifiers and a different range of outputs, the min-max score normalization method was applied to scale the results of each technique into the range between 0 and 1. Based upon prior research, two fusion methods were utilized: simple sum and matcher weighting [11], [29]. For the Simple Sum fusion, the raw score of each individual technique were simple added and rescaled into the 0 to 1 range. For the Matcher Weighting approach, weights are assigned to the individual matchers based on their individual EER. The weights are inversely proportional to the corresponding errors; the weights for less EER are higher than those of with a high EER.

3.2 Experiment results

The results of using individual biometrics and the multimodal approach are shown in Table 2. The results illustrated an average of all the users' EERs by using a single optimized neural network. The results showed that the individual techniques can be used to discriminate users with relatively low error rates for a good proportion of participants. However, further analysis showed that the individual user is able to achieve a better overall EER when each user is permitted to use a different network configuration. By using individually optimized network configurations for individual user, the overall performance was an EER of 8.9%. Behavioral profiling demonstrated the best individual performance using a single network configuration, with keystroke dynamics being the worst performer.

A further analysis of individual performances raises a number of interesting points. Foremost, that the best-case EERs are extremely good. However, it is noticeable that there are some users that experience very high error rates, reiterating the importance of multimodal approaches.

Table 2. Experiment results for text-based authentication

		Equal Error Rate (EER)%		
		Average	Best Case	Worst Case
Linguistic Profiling (LP)		12.8	0.0	40.0
Behavioral Profiling (BP)		9.2	0.0	50.0
Keystroke Dynamics (KA)		20.8	0.0	50.7
Fusion by Sum				
	Multimodal (LP+BP)	5.5	0.0	30.6
	Multimodal (KA+BP)	6.2	0.0	20.0
	Multimodal (LP+KA)	11.2	0.0	45.0
	All techniques	4.4	0.0	18.1
Fusion by Matcher Weighting				
	Multimodal (LP+BP)	3.6	0.0	20.0
	Multimodal (KA+BP)	5.3	0.0	20.2
	Multimodal (LP+KA)	8.5	0.0	44.7
	All techniques	3.3	0.0	19.3

As seen in the Table 2, both of the two fusion methods lead to better performance than any of the individual classifiers. Generally, the Matcher Weighting technique outperforms simple sum method. Whilst the results show that on average the use of more modalities leads to a better performance, this is not reflected within the individual user results. On occasions, it was noticed that users performed better when using two inputs (typically LP+BP) rather than three. Therefore in an operational environment case must be taken on selecting the most appropriate classifier. Examining the individual worst-case performance, it can be seen that the multimodal models have significantly improved upon the error rates – further supporting the use of multimodal approaches.

4 A novel framework for active authentication

The concept of Transparent Authentication System (TAS) on mobile devices was first proposed in 2002 [30]. The framework utilizes a mixture of biometric techniques to verify a mobile user's identity in a continuous and transparent manner. The framework is able to:

- to increase the authentication security beyond that offered by the password based approach;
- to provide transparent non-intrusive authentication for the user (rather than intrusive) to maximize user convenience;
- to provide continuous verification of the user, ensuring that the protection can be maintained throughout the duration of the device usage;
- to provide an authentication architecture that automatically works on all mobile devices regardless of hardware configuration, processing capability and network connectivity.

A number of process engines and a security manager have been devised to achieve these objectives (as demonstrate in Fig.1). A detailed description of these processes is presented in the following sections.

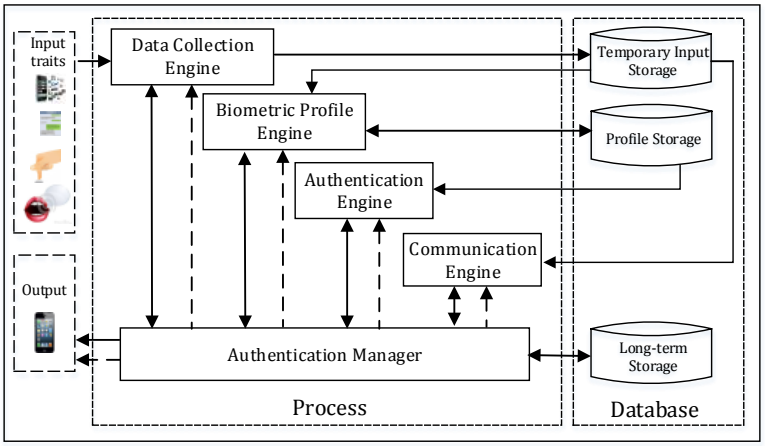

Fig. 1. Text-based multimodal framework

4.1 Processing engines

The primary role of the Data Collection Engine is to capture a user's input text. When a user utilizes a text-based application on the mobile, information about the user's typing, message writing style and the application usage are automatically collected by the Data Collection Engine and transformed into various biometric input samples. The captured input samples are then stored in the Temporary Input Storage to be used further in the authentication process by the authentication engine.

The main duty of the biometric profiling engine is to generate the various biometric profile templates by using the combination of the user's historical data and a number of template generation algorithms. The generated biometric templates will be stored in the Profile Storage and will be used in the verification process.

The main functionality of the Authentication Engine is to perform the user authentication process. The Authentication Engine has the ability to perform authentication for every permutation of inputs to ensure that authentication can be performed even if all of the three biometric samples are not presented (e.g. location may be not be able to be determined). When a verification process is required by the Authentication Manager, the Authentication Engine compares the input samples with the biometric templates to determine the legitimacy of the user. Once the verification process is completed, the verification result is appropriately processed by the Authentication Manager. If the verification result indicates the sample(s) came from authorized user, the sample(s) will be stored within the Profile Storage to be used for profile (re)generation; otherwise it will be deleted. A multibiometric authentication technique may produce a verification result that accepts the samples as coming from the authorized user even though the sample from one individual technique might be rejected as

coming from an imposter. Since the overall decision was that the sample comes from the authorized user, the failed samples are deemed to be, in fact, from the authorized user and incorrectly failed. As such, these samples are added to the profile and are not deleted. In this way, the template re-training process can produce a more accurate profile that could provide better performance. This process overcomes a fundamental issue with biometric template re-training and ensuring the correct inclusion of relevant samples.

The framework can operate in both standalone and distributed modes to allow the framework to be useful for non-wireless and wireless devices. If the framework operates in client-server mode, the communication engine works as a bridge between the capture device and the comprehensive framework. When the framework operates in a standalone mode and the device is locked down, the communication engine sends a code to the user which they can use to unlock their device.

4.2 Security manager

The Authentication Manager is the central controller of the framework and provides the "intelligence". The key task of the Authentication Manager is to monitor the security level and make authentication decisions when the user requests access to an application. It is the responsibility of the Authentication Manager to handle the security and user convenience trade-off. In order to achieve this, the Authentication Manager utilizes two processing algorithms: the System Security (SS) Level Automatic Update Algorithm and the Application Request Algorithm to manage the balance between the security of the mobile device and user convenience. These processes have been designed based upon a well-known study [24].

The SS level is a sliding numerical value in the range of 0 and +5 with 0 indicating a low security level and +5 indicates a high security level[1]. The SS level changes depending upon the outcomes of the authentication processes and the time that has elapsed between authentication requests. In this proposed framework, each application will have its own security level. The high value application will have a high security level and a normal application will have a low security level. This can be achieved either manually by the user or automatically by the system, using a database stored in the Long-term Storage. Prior research has investigated simple mechanisms by which these risk-based evaluations for applications can be made [30].

The Authentication Manager utilizes the SS Level Automatic Update Algorithm in order to periodically update the SS level based on the results of authentication decisions based upon the user's input samples. The Authentication Manager periodically sends an authentication request to the Authentication Engine in order to update the SS level. The time interval in which the authentication should be requested depends upon the user's preference (i.e. every 5 minutes). Initially, the Authentication Manager requires the Authentication Engine to perform authentication using the best set of the user's input samples (i.e. utilizes the classifier with the lowest EER that samples exist

[1] The boundaries defined on the numerical scale are only provided as a suggestion. In practice, these values maybe redefined.

for) from the last x minutes (i.e. 5 minutes). In a case where no user's input data is presented, the Authentication Manager maintains the SS level at its latest updated value. However, if the Authentication Engine responds with a pass then the Authentication Manager updates the SS level and subsequently reverts back to monitoring mode. If not, the Authentication Manager decreases the SS level and sends an authentication request again by using the next best set of user's input samples. The Authentication Manager will try three times to send an authentication request, every time with the next best available sample being employed. The Authentication Manager updates the SS level based upon the authentication result. The SS value is increased or decreased based on the type of sample used. For example, a sample using the keystroke dynamics technique will have an increment/decrement value of 0.5; a sample which contains both linguistic profiling and behavior profiling will have an increment/decrement value of 2. This numbers are based on the performance of the technique or combination of techniques. In scenarios where the updated SS level is less than 0, the Authentication Manager will set the SS level back to 0, meaning that the user will be able to access only the applications that do not required security. The process gives bias toward the user as they are given three non-intrusive chances to authenticate correctly and no intrusive authentication requests. This enables the system to minimize inconvenience to its user. Should the user attempt to access applications that require a SS level greater than the current SS level, the Authentication Manager will utilize the Application Request Algorithm to check the legitimacy of the user as shows in Fig.2.

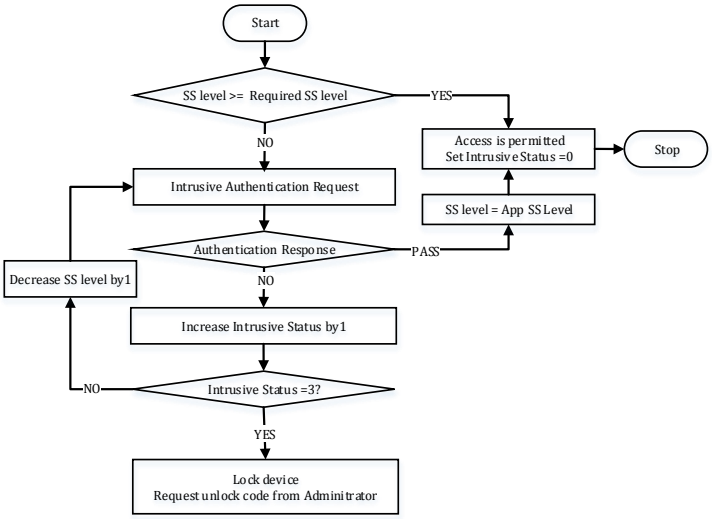

Fig. 2. Application Request Algorithm

The current SS level of the user is compared with the security level of the requested application. If the level is equal to or greater than the security level of the required application, the user can automatically access the application. Otherwise the user will be asked intrusively to authenticate. If the authentication response to this intrusive

request fails to pass, the device is locked. Otherwise, the level of the user will be updated to the security level of the requested application and access will be granted.

5 Evaluation

To examine the effectiveness of the framework in providing security and user convenience, the proposed framework was evaluated through a simulation. The simulation process involves implementing a virtual user and applying the SS Level Automatic Update Algorithm and the Application Request Algorithms.

To evaluate the performance of the security mechanisms to an authorized user, three different usage levels (infrequent, moderate and frequent) will be investigated - as the level of usage will have a direct impact on the availability of biometric samples and thus the capability of the system to maintain the security level. The use of the mobile device is simulated using a flow of timeslots. Each time slot can be seen as a minute in real life. Within each time slot the user can do one of two actions, or both: provide an input sample (thus simulating a text-based entry) or the use an app. Within each timeslot, the probability for the user to provide an input sample or accessing an application will set to 0.05, 0.15 and 0.50 in order to simulate an infrequent, moderate and frequent user respectively. There are 6 different types of application that can be chosen by the user (reflecting the possible security levels of an application from 0 to 5). Each type of application has the same probability of being accessed. Similarly, there are 7 different non-intrusive techniques (refer to Table 2). Given that within a time slot the user provides an input sample, each type of technique has the same probability of occurring.

All non-intrusive techniques are evaluated based upon the EER of each authentication technique as demonstrated in the experimental result section. This means that, when the system evaluates a sample, there is a probability (equal to the EER of the technique) that an authorized user will be rejected or an imposter will be authorized. With regards to the intrusive authentication requests, the probability of an authorized user and impostor being rejected and accepted respectively is set to 0.03. This approach to the methodology removes any bias and provides for a randomly generated dataset with a mix of samples, performances and application requests across three usage scenarios. To further remove any bias that would exist from a single run of the simulation, the simulation is repeated.

The security system will work as described in the Security Manager session. The SS will be updated every 10 minutes. If the mobile device is not used for 10 minutes consecutively, the SS will be decreased by 0.05 for every following minute, until the system is used again. The simulation simulated the use of the mobile phone for 12 hours or 720 minutes.

In order to examine the ability of the system security to prevent an imposter from using the mobile device, two scenarios were simulated: an imposter using a mobile device at the initial state (SS =0) and the imposter using a mobile device starting from a high level of security (SS=5). This can simulate an imposter taking control a mobile device which has just been used by the authorized user.

5.1 Simulation results

The result for all scenarios is represented using the average of running the simulation 10 times. The simulation results for an infrequent, moderate and frequent authorized user are presented in Table3.

Table 3. Simulation results for different types of authorized user

	Infrequent User		Moderate User		Frequent User	
App Level	**#App Request**	**#Intrusive Request**	**#App Request**	**# Intrusive Request**	**#App Request**	**# Intrusive Request**
5	7.2	4.2	16.2	1.5	60.0	1.50
4	5.1	0.4	17.9	0.5	60.1	0.50
3	7.3	0.3	20.0	0.2	61.3	0.30
2	5.9	0.2	16.8	0.3	59.6	0.10
1	6.0	0.0	19.5	0.2	55.2	0.00
0	6.7	0.0	19.3	0.0	57.6	0.00
Total	38.2	5.1	109.7	2.7	353.8	2.40

Based upon the simulation results, it can be shown that the security system can provide a high level of security whilst minimizing user inconvenience in all three scenarios. Analysing the proportion between intrusive authentication request and application access permits an insight into how often the user experiences an intrusive authentication request. Ideally, this proportion would be zero meaning that the user would not be required to perform an intrusive authentication request when they access an application. In our simulation these values are 13%, 2% and 0.6%, for the infrequent, moderate and frequent user respectively. The infrequent user experiences a higher intrusive request because it will probabilistically have fewer samples in the system and the system decreases the SS level if the device is not used for 10 consecutive minutes. Therefore, when this user want to access an application, it is more likely that its SS will not be sufficient to be granted immediate access. Throughout the complete 720 minute simulation the device was never incorrectly blocked for the authorized user. Further analysis of the results demonstrates for a level 5 app (which is arguably sensitive enough to warrant authentication of the user), this transparent approach results in a 97.5% reduction in intrusive authentication requests (for a frequent user).

The simulation results of the imposter scenarios showed that the security system blocks the imposter from using the mobile device after few minutes in both cases (as illustrated in Table 4). The reasons for this is that when the imposter tried to access an application that required a security level greater than 0, the system requested the imposter to authenticate themselves using an intrusive technique three times. There is a really small chance for the imposter to successfully authenticate, so after three requests the device will be blocked. As expected, the system will take more time to block the device if the imposter starts using the device when the SS is high.

Table 4. Simulation results for imposter user start using device at SS=0 and 5

	Device at SS= 0		Device at SS= 5	
App Level	**#App Request**	**# Intrusive Request**	**#App Request**	**# Intrusive Request**
5	0.6	0.4	1.0	0.2
4	0.3	0.1	1.8	0.6
3	0.4	0.4	1.5	0.0
2	0.4	0.2	1.3	0.1
1	0.3	0.1	1.9	0.1
0	0.4	0.0	1.0	0.0
Total	2.4	1.2	8.5	1.0
Time In Use	5.0 minutes		14.7 minutes	

5.2 Discussion

The simulations show how the proposed framework can provide a good compromise between improving the level of security provided and without increasing the user convenience. Indeed, it can be argued that user convenience under this model is also significantly improved over existing approaches. However, further investigations are required in order to better examine the values of the parameters. For example, it seems clear that the verification time does play an important role of providing security and user convenience. By regularly authenticating the user, the user will suffer more intrusive authentication requests but the system will be able to recognize an imposter in a relatively short period of time. On the other hand, users will find the device more convenient to use with longer time periods between user authentications but the system will take longer to recognize an imposter and lock down the system. In our simulation the verification time was 10 minutes. However, this may not be the optimum compromise between convenience and security.

Similarly, decreasing SS level was not examined, but it is expected to play a relevant role in the system. The infrequent user will experience less challenges from the intrusive authentication technique when the time period of the degradation function gets longer. However, the imposter will have more chance of accessing a high level application in cases where the device was initially left with a high level SS. In this simulation, a linear function is used to decrease the SS level but it is suggested that the function for degrading the SS level should be implemented using an exponential function as it decrease slowly at first and then more rapidly.

6 Conclusions & Future Work

The first part of this paper presented a feasibility study that demonstrated the ability of utilizing text-based entry to authenticate users. The use multimodal biometrics, specifically the combination of linguistic profiling, behavior profiling and keystroke dynamics showed an excellent level of recognition performance, validating the feasi-

bility that multimodal text-based has the ability to authenticate user on mobile devices.

The novel multimodal authentication framework subsequently presented to support text-based biometrics was designed to add additional security to a mobile handset, providing transparent and continuous authentication. The system is designed using a variety of single and multimodal biometric techniques without any additional hardware. The users can benefit from the framework in terms of both device security and convenience of use. By setting various security requirement levels for different applications/services based upon their risk, the framework is capable of controlling the impact on each application/service. The simulation results clearly showed that the proposed authentication framework is able to provide continuous and transparent authentication to protect mobile devices.

Future work will focus upon the development of a more representative and larger biometric corpus from which to further examine the level of recognition performance that can be achieved. To accompany this work, an operational prototype will also be developed to enable an end-user evaluation to be undertaken so that user acceptance and operational performance can be established.

7 References

1. Ericsson.: Traffic and market report on the pulse of the networked society, http://www.ericsson.com/res/docs/2012/traffic_and_ market_report_june_2012.pdf
2. Kaspersky Lab.: European Users Mobile Behaviour and Awareness of MobileThreats, http://www.kaspersky.com/news?id=207576289
3. Dimensional Research.: The impact of mobile devices on information security: A survey of IT professionals, http://www.checkpoint.com/downloads/products/check-point-mobile-security-survey-report.pdf
4. McAfee.: McAfee Reveals Consumers Fail To Protect Their Mobile Devices, http://www.mcafee.com/us/about/news/2013/q1/ 20130224-01.aspx
5. Clarke, N. and Furnell, S.M.: Authentication of users on mobile telephones – A survey of attitudes and practices, Computer & Security, vol.24, no. 7, pp519-527, (2005)
6. Karatzouni S., Clarke, N. and Furnell, M.: Utilising Biometric for transparent user authentication on mobile devices. In: 2nd Internet Technologies and Applications, pp.549-557 (2007)
7. Li, Fudong, Nathan Clarke, Maria Papadaki, and Paul Dowland.: Behaviour Profiling for Transparent Authentication for Mobile Devices. In *Proceedings of the 10th European Conference on Information Warfare (ECIW), Tallinn, Estonia*, pp. 307-314. (2011)
8. Sim, T., Zhang, S., Janakiraman, R., & Kumar, S.: Continuous verification using multimodal biometrics. In: Pattern Analysis and Machine Intelligence, vol 29, no 4, pp. 687-700. (2007)
9. Kittler, J.,Matas, J., Jonsson, K. and Ramos Sanchez, M. U.: Combining Evidence in Personal Identity Verification Systems. Pattern Recognition Letters, vol.18, pp.845-852 (1997)
10. Poh, N. and Korczak, J.: Hybrid Biometric Authentication System Using Face and Voice Features. Lecture Notes in Computer Science, vol.2091/2001, pp. 348-353 (2001)
11. Snelick, R., Uludag, U., Mink, A., Indovina, M., and Jain, A.K.: Large-Scale Evaluation of Multimodal Biometric Authentication Using State-of-the-Art Systems. IEEE Transactions on pattern analysis and machine intelligence, Vol.27, no. 4 pp.450-455, (2005)

12. Cmo Council, http://www.fastcompany.com/3010237/bottom-line/texting-is-the-new-email-does-your-company-do-it-right
13. ComputerWeekly,http://www.computerweekly.com/news/2240205200/Apple-adopts-hands-off-approach-to-iPhone-fingerprint-scanner
14. MIT Technology Review, http://www.technologyreview.com/news/425805/new-google-smart-phone-recognizes-your-face/
15. Rudman, J.: The state of authorship attribution studies: Some problems and solutions. Computers and the Humanities, 31, 351-365. (1998)
16. Halteren, V. H.: Linguistic Profiling for Author Recognition and Verification, In: 42nd Annual Meeting on Association for Computational Linguistics (ACL 04), Association for Computational Linguistics, Morristown, NJ, USA, (2004)
17. Zheng R., Li, J., Chen, H., and Huang Z.: A Framework for Authorship Identification of Online Messages: Writing-Style Features and Classification Techniques. Journal of the American Society for Information Science and Technology, vol. 53, pp. 378-393. (2006)
18. Boukerche, A., Nitare,M.S.M.A.: Behavior-based intrusion detection in mobile phone systems. J. Parallel Distrib. Comput. **62**(9),1476–1490 (2002)
19. Damopoulos, D. Menesidou, S. Kambourakis, Papadaki, M. Clarke, N. Gritzalis, S. Evaluation of Anomaly-Based IDS for Mobile Devices Using Machine Learning Classifiers, Security and Communication Networks, Vol. 5, No. 1, pp. 3-14, 2012, Wiley
20. Buschkes, R., Kesdogan, D., Reichl, P.: How to increase security in mobile networks by anomaly detection. In: Proceedings of the 14th Annual Computer Security Applications Conference, pp. 3–12 (1998)
21. Hall, J.,Barbeau,M., Kranakis, E.: Anomaly based intrusion detection using mobility profiles of public transportation users. In: Proceeding of IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, vol. 2, pp. 17–24 (2005)
22. Biopassword.: the keystroke dynamics approach, http://www.biopassword.com/bp2/welcome.asp.
23. Behaviosec, http://www.behaviosec.com/products/enterprise/
24. Clarke, N. and Furnell, S.M.: Authenticating Mobile Phone Users Using Keystroke Analysis, International Journal of Information Security, ISSN: 1615-5262, pp.1-14. (2006)
25. Indovina, M., Uludag, U., Snelick, R., Mink, A., & Jain, A.: Multimodal biometric authentication methods: a COTS approach. *Proc. MMUA*, 99-106. (2003)
26. Eagle, N., Pentland, A., Lazer, D.: inferring Social Network Structure using Mobile Phone Data, Proceeding of National Academy of Sciences (PNAS),vol.106, pp.15274-1578, (2009)
27. Ashbourne, J.: Biometric, Advanced identity verification. The complete guide. Springer, (2000)
28. Jain, A. K. Nandakumar, K. and Ross, A.: Score normalization in multimodal biometric systems. Pattern Recognition, vol. 38, no. 12, pp.2270-2285, Dec. (2005)
29. Clarke, N., Furnell, S.M. and Reynolds P.L.: Biometric Authenticating for Mobile Devices. In: 3rd Australian Information Warfare and Security Conference, Western Australia, (2002)
30. Ledermuller, T. and Clarke, N.L.: Risk assessment for mobile devices. In: 8th International Conference Privacy and Security in Digital Business, TrustBus, pp. 210–221 (2011)